

RECEIVED

OCT 03 2001

TECH CENTER 1600/2900

1652 Page 1 of 6

#7
Dmt
10-4-01

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/619,032

DATE: 09/27/2001

TIME: 15:42:46

Input Set : N:\Crf3\RULE60\09619032.txt

Output Set: N:\CRF3\09272001\I619032.raw

SEQUENCE LISTING

C--> 5 (1) GENERAL INFORMATION:

7 (i) APPLICANT: Murphy, Dennis
8 Reid, John

C--> 10 (ii) TITLE OF INVENTION: ALPHA-GALACTOSIDASE

12 (iii) NUMBER OF SEQUENCES: 4

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Fish & Richardson, P.C.
16 (B) STREET: 4225 Executive Square, Suite 1400
17 (C) CITY: La Jolla
18 (D) STATE: CA
19 (E) COUNTRY: US
20 (F) ZIP: 92037

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Diskette
24 (B) COMPUTER: IBM Compatible
25 (C) OPERATING SYSTEM: Windows95
26 (D) SOFTWARE: FastSEQ for Windows Version 2.0

28 (vi) CURRENT APPLICATION DATA:

C--> 29 (A) APPLICATION NUMBER: US/09/619,032
C--> 30 (B) FILING DATE: 19-Jul-2000

31 (C) CLASSIFICATION:

33 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: 09/407,806
35 (B) FILING DATE:

37 (viii) ATTORNEY/AGENT INFORMATION:

38 (A) NAME: Haile, Ph.D., Lisa A.
39 (B) REGISTRATION NUMBER: 38,347
40 (C) REFERENCE/DOCKET NUMBER: 09010/004001

42 (ix) TELECOMMUNICATION INFORMATION:

43 (A) TELEPHONE: 619-678-5070
44 (B) TELEFAX: 619-68-5099
45 (C) TELEX:

47 (2) INFORMATION FOR SEQ ID NO: 1:

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 52 base pairs
51 (B) TYPE: nucleic acid
52 (C) STRANDEDNESS: single
53 (D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: cDNA

57 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

59 CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAGAGCG CTCGTCTTTC AC

52

61 (2) INFORMATION FOR SEQ ID NO: 2:

63 (i) SEQUENCE CHARACTERISTICS:

64 (A) LENGTH: 31 base pairs
65 (B) TYPE: nucleic acid
66 (C) STRANDEDNESS: single

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/619,032

DATE: 09/27/2001
TIME: 15:42:46

Input Set : N:\Cr3\RULE60\09619032.txt
Output Set: N:\CRF3\09272001\I619032.raw

```

67      (D) TOPOLOGY: linear
69      (ii) MOLECULE TYPE: cDNA
71      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:                               31
73      CGGAAGATCT AGGTTCCCCA TTTTCACCCC T
75 (2) INFORMATION FOR SEQ ID NO: 3:
77      (i) SEQUENCE CHARACTERISTICS:
78          (A) LENGTH: 1041 base pairs
79          (B) TYPE: nucleic acid
80          (C) STRANDEDNESS: single
81          (D) TOPOLOGY: linear
83      (ix) FEATURE:
84          (A) NAME/KEY: Coding Sequence
85          (B) LOCATION: 1...1038
86          (D) OTHER INFORMATION:
88      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:                               48
90      TTG AGA GCG CTC GTC TTT CAC GGC AAC CTC CAG TAT GCC GAA ATC CCA
91      Leu Arg Ala Leu Val Phe His Gly Asn Leu Gln Tyr Ala Glu Ile Pro
92      1          5          10          15
94      AAG AGC GAA CCA AAG GTC ATA GAG AAG GCA TAC ATC CCA GTC ATC GAG       96
95      Lys Ser Glu Pro Lys Val Ile Glu Lys Ala Tyr Ile Pro Val Ile Glu
96      20          25          30
98      ACA CTG ATT AAA GAA GAA CCT TTT GGG CTC AAC ATA ACG GGC TAT ACC       144
99      Thr Leu Ile Lys Glu Glu Pro Phe Gly Leu Asn Ile Thr Gly Tyr Thr
100     35          40          45
102     TTA AAG TTC CTC CCG AAG GAT ATT ATA CTC GTT AAA GGG GGC ATC GCG       192
103     Leu Lys Phe Leu Pro Lys Asp Ile Ile Leu Val Lys Gly Gly Ile Ala
104     50          55          60
106     AGT GAC CTG ATA GAG ATA ATC GGA ACG AGC TAC ACG GCA ATA CTC CCC       240
107     Ser Asp Leu Ile Glu Ile Ile Gly Thr Ser Tyr Thr Ala Ile Leu Pro
108     65          70          75          80
110     CTC CTG CCG CTT AGC AGA GTA GAA GCA CAA GTT CAG AGA GAT AGG GTT       288
111     Leu Leu Pro Leu Ser Arg Val Glu Ala Gln Val Gln Arg Asp Arg Val
112     85          90          95
114     AAG GAA GAG CTC TTC GAG GTT TCT CCA AAG GGA TTC TGG CTG CCA GAG       336
115     Lys Glu Glu Leu Phe Glu Val Ser Pro Lys Gly Phe Trp Leu Pro Glu
116     100         105         110
118     CTC GCC GAC CCG ATA ATC CCT GCC ATA CTG AAG GAC AAC GGT TAT GAG       384
119     Leu Ala Asp Pro Ile Ile Pro Ala Ile Leu Lys Asp Asn Gly Tyr Glu
120     115         120         125
122     TAT CTA TTC GCC GAC GAG GCG ATG CTT TTC TCA GCT CAT CTC AAC TCG       432
123     Tyr Leu Phe Ala Asp Glu Ala Met Leu Phe Ser Ala His Leu Asn Ser
124     130         135         140
126     GCG ATA AAG CCA ATT AAA CCG CTC CCA CAC CTT ATA AAG GCC CAA AGG       480
127     Ala Ile Lys Pro Ile Lys Pro Leu Pro His Leu Ile Lys Ala Gln Arg
128     145         150         155
130     GAA AAG CGC TTT AGG TAC ATC AGC TAT CTC CTT CTC AGG GAG CTT AGG       528
131     Glu Lys Arg Phe Arg Tyr Ile Ser Tyr Leu Leu Leu Arg Glu Leu Arg
132     165         170         175
134     AAG GCG ATA AAG CTC GTT TTT GAA GGT AAG GTA ACG CTA AAG GTC AAA       576

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/619,032

DATE: 09/27/2001
TIME: 15:42:46

Input Set : N:\Crif3\RULE60\09619032.txt
Output Set: N:\CRF3\09272001\I619032.raw

```

135 Lys Ala Ile Lys Leu Val Phe Glu Gly Lys Val Thr Leu Lys Val Lys
136          180          185          190
138 GAC ATC GAA GCC GTA CCC GTT TGG GTG GCC GTG AAC ACG GCT GTA ATG 624
139 Asp Ile Glu Ala Val Pro Val Trp Val Ala Val Asn Thr Ala Val Met
140          195          200          205
142 CTC ATC GGA AGG CTT CCT CTT ATG AAT CCT AAG AAA GTG GCG AGC TGG 672
143 Leu Ile Gly Arg Leu Pro Leu Met Asn Pro Lys Lys Val Ala Ser Trp
144          210          215          220
146 ATA GAG GAC AAG AAC ATT CTT CTA TAC GGC ACC GAT ATA GAG TTC ATT 720
147 Ile Glu Asp Lys Asn Ile Leu Leu Tyr Gly Thr Asp Ile Glu Phe Ile
148          225          230          235
150 GGC TAT AGG GAC ATT GCA GGC AGA ATG AGT GTT GAG GGA TTA TTA GAG 768
151 Gly Tyr Arg Asp Ile Ala Gly Arg Met Ser Val Glu Gly Leu Leu Glu
152          245          250          255
154 GTT ATA GAC GAG CTC AAC TCG GAA CTG TGC CCC TCA GAG CTG AAG CAC 816
155 Val Ile Asp Glu Leu Asn Ser Glu Leu Cys Pro Ser Glu Leu Lys His
156          260          265          270
158 AGT GGA AGG GAG CTC TAC TTA CGG ACT TCG AGT TGG GCA GAT AAG AGC 864
159 Ser Gly Arg Glu Leu Tyr Leu Arg Thr Ser Ser Trp Ala Asp Lys Ser
160          275          280          285
162 TTG AGG ATA TGG AGA GAG GAC GAA GGG AAC GCA AGA CTT AAT ATG CTG 912
163 Leu Arg Ile Trp Arg Glu Asp Glu Gly Asn Ala Arg Leu Asn Met Leu
164          290          295          300
166 TAC AAT ATG AGG GGC GAA CTC GCC TTT TTA GCC GAG AAC AGC GAT GCA 960
167 Tyr Asn Met Arg Gly Glu Leu Ala Phe Leu Ala Glu Asn Ser Asp Ala
168          305          310          315
170 AGG GGA TGG CCC CTC CCT GAG AGG AGG CTG GAT GCC TTC CGG GCG ATA 1008
171 Arg Gly Trp Pro Leu Pro Glu Arg Arg Leu Asp Ala Phe Arg Ala Ile
172          325          330          335
174 TAT AAC GAT TGG AGG GGT AAT GGG GAA CCT TAG 1041
175 Tyr Asn Asp Trp Arg Gly Asn Gly Glu Pro
176          340          345
179 (2) INFORMATION FOR SEQ ID NO: 4:
181 (i) SEQUENCE CHARACTERISTICS:
182 (A) LENGTH: 346 amino acids
183 (B) TYPE: amino acid
184 (D) TOPOLOGY: linear
186 (ii) MOLECULE TYPE: protein
188 (v) FRAGMENT TYPE: internal
190 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
192 Leu Arg Ala Leu Val Phe His Gly Asn Leu Gln Tyr Ala Glu Ile Pro
193 1 5 10 15
194 Lys Ser Glu Pro Lys Val Ile Glu Lys Ala Tyr Ile Pro Val Ile Glu
195 20 25 30
196 Thr Leu Ile Lys Glu Glu Pro Phe Gly Leu Asn Ile Thr Gly Tyr Thr
197 35 40 45
198 Leu Lys Phe Leu Pro Lys Asp Ile Ile Leu Val Lys Gly Gly Ile Ala
199 50 55 60
200 Ser Asp Leu Ile Glu Ile Ile Gly Thr Ser Tyr Thr Ala Ile Leu Pro

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/619,032

DATE: 09/27/2001

TIME: 15:42:46

Input Set : N:\Crif3\RULE60\09619032.txt
 Output Set: N:\CRF3\09272001\I619032.raw

201	65				70					75				80
202	Leu	Leu	Pro	Leu	Ser	Arg	Val	Glu	Ala	Gln	Val	Gln	Arg	Asp
203														Arg
204	Lys	Glu	Glu	Leu	Phe	Glu	Val	Ser	Pro	Lys	Gly	Phe	Trp	Leu
205														Pro
206	Leu	Ala	Asp	Pro	Ile	Ile	Pro	Ala	Ile	Leu	Lys	Asp	Asn	Gly
207														Tyr
208	Tyr	Leu	Phe	Ala	Asp	Glu	Ala	Met	Leu	Phe	Ser	Ala	His	Leu
209														Asn
210	Ala	Ile	Lys	Pro	Ile	Lys	Pro	Leu	Pro	His	Leu	Ile	Lys	Ala
211														Gln
212	Glu	Lys	Arg	Phe	Arg	Tyr	Ile	Ser	Tyr	Leu	Leu	Leu	Arg	Glu
213														Leu
214	Lys	Ala	Ile	Lys	Leu	Val	Phe	Glu	Gly	Lys	Val	Thr	Leu	Lys
215														Val
216	Asp	Ile	Glu	Ala	Val	Pro	Val	Trp	Val	Ala	Val	Asn	Thr	Ala
217														Val
218	Leu	Ile	Gly	Arg	Leu	Pro	Leu	Met	Asn	Pro	Lys	Lys	Val	Ala
219														Ser
220	Ile	Glu	Asp	Lys	Asn	Ile	Leu	Leu	Tyr	Gly	Thr	Asp	Ile	Glu
221														Phe
222	Gly	Tyr	Arg	Asp	Ile	Ala	Gly	Arg	Met	Ser	Val	Glu	Gly	Leu
223														Leu
224	Val	Ile	Asp	Glu	Leu	Asn	Ser	Glu	Leu	Cys	Pro	Ser	Glu	Leu
225														Lys
226	Ser	Gly	Arg	Glu	Leu	Tyr	Leu	Arg	Thr	Ser	Ser	Trp	Ala	Asp
227														Lys
228	Leu	Arg	Ile	Trp	Arg	Glu	Asp	Glu	Gly	Asn	Ala	Arg	Leu	Asn
229														Met
230	Tyr	Asn	Met	Arg	Gly	Glu	Leu	Ala	Phe	Leu	Ala	Glu	Asn	Ser
231														Asp
232	Arg	Gly	Trp	Pro	Leu	Pro	Glu	Arg	Arg	Leu	Asp	Ala	Phe	Arg
233														Ala
234	Tyr	Asn	Asp	Trp	Arg	Gly	Asn	Gly	Glu	Pro				Ile
235														

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/619,032

DATE: 09/27/2001
TIME: 15:42:47

Input Set : N:\Crf3\RULE60\09619032.txt
Output Set: N:\CRF3\09272001\I619032.raw

L:5 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]
L:10 M:220 C: Keyword misspelled or invalid format, [(ii) TITLE OF INVENTION:]
L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]